Report of the committee on nomenclature

Prepared by:
E. R. Giblett

Committee members:

E. R. Giblett, Seattle (chairman)
H. Harris, London
P. Meera Khan, Leiden
E. W. Lovrien, Portland
W. J. Mellman, Philadelphia
C. W. H. Partridge, New Haven
T. B. Shows, Buffalo

The guidelines outlined below are based on the decisions made at an interim meeting of the committee held on April 3-4, 1975. These guidelines, including a list of suggested names for enzyme loci, were approved during a plenary session of the Gene Mapping Workshop in Baltimore, 1975.

Consideration was confined to enzyme nomenclature because the blood group antigens have been dealt with in the standard text by Race and Sanger, while participants in the Histocompatibility Testing workshops have handled the terminology problems in that area. In general, the guidelines for naming enzymes can also be applied to the plasma proteins.

Guidelines for Genetic Nomenclature of Human Enzymes
(as already reported)

Addendum: Two enzymes not listed above, adenosine kinase and β-glucuronidase, were assigned to chromosomal loci at the Baltimore meeting. It is proposed that their respective loci be designated ADK and βGUS.

REPORT OF NOMENCLATURE COMMITTEE

On April 3-4, 1975, a meeting was held in Philadelphia by the following appointed and ad hoc members: Harry Harris, Meera Khan, Everett Lovrien, William Mellman, Chester Partridge and Thomas Shows, with Eloise Giblett as chairman.

At the outset, it was decided that this committee should confine itself to enzyme nomenclature, and that the guidelines for naming enzymes could probably be applied to plasma proteins, as well. The nomenclature used for blood group antigens has been dealt with extensively in the standard text by Race and Sanger, while the participants of the HL-A workshops have worked out their own terminology for the histocompatibility loci and their products.

Dr. Harris had already developed a set of guidelines before the committee meeting, and it was used as the basis of our deliberations. The consensus emerging from these discussions is presented in the following paragraphs.

Guidelines for Genetic Nomenclature of Human Enzymes

The essentials of a satisfactory terminology are that it should be precise and unambiguous, it should clearly distinguish between genotypes and phenotypes, and as far as possible, the symbols used should readily identify the particular enzyme. In addition, it should be sufficiently flexible to permit incorporation of some unusual symbols used to designate certain enzymes in the original papers which have subsequently been widely adopted and embedded in the literature. It should also be capable of incorporating new discoveries as they are reported.

The following general scheme appears to meet these requirements and is reasonably convenient in practice.

I. Genotypes

Genotypic symbols, i.e. for loci or alleles, are italicised (or underlined in typescript) to distinguish them clearly from symbols used to designate phenotypes, which are not italicised or underlined.

A. Loci

1. Loci are designated by letters, either all capitalized (preferred) or just the first letter. Usually two or three letters will suffice, but sometimes four or even five may be required.

   Examples: ADA for adenosine deaminase
   UMPK for uridine monophosphate kinase
   Gd for glucose-6-phosphate dehydrogenase

   (Obviously, when lower case letters are used to designate one locus, it is undesirable to use the same letters but in capitals (e.g. GD) to designate another locus.) The letters chosen for locus names are preferably based on the recommended name given by the Enzyme Commission on Nomenclature. However, this is sometimes inconvenient or confusing because of past usage. Thus, GOT is preferred for glutamic-oxaloacetic transaminase, although the E.C. recommended name is aspartate aminotransferase. In some cases, Greek letters are also needed for clarity. Example: αGAL for α-galactosidase to distinguish it from β-galactosidase (βGAL).

2. There are often two or more loci coding for different polypeptide chains which are contained in separate enzyme proteins having very similar or identical catalytic properties. Such loci are best differentiated by appropriate subscripts.

Examples:

$\underline{PGM_1}$, $\underline{PGM_2}$ and $\underline{PGM_3}$ for the three phosphoglucomutase loci

$\underline{ADH_1}$, $\underline{ADH_2}$ and $\underline{ADH_3}$ for the three alcohol dehydrogenase loci

Numerical subscripts are often most convenient. However, sometimes, because of past usage or easy identification, letters are preferred to avoid confusion.

Examples:

$\underline{LDH_A}$, $\underline{LDH_B}$ and $\underline{LDH_C}$ for the three lactate dehydrogenase loci

$\underline{PGAM_M}$ and $\underline{PGAM_B}$ for the two phosphoglycerate mutase loci which are active in muscle and brain, respectively.

Some enzymes occur in a so-called soluble (or supernatant or cytosol) form and also in a mitochondrial form, with the two forms being catalytically similar but coded at separate loci. In such cases, the use of S and M as subscripts may be less confusing than numerical or alphabetical designations.

Example:

$\underline{GOT_S}$ and $\underline{GOT_M}$ for the soluble and mitochondrial forms of glutamic-oxaloacetic transaminase.

B. Alleles

Different alleles at the same locus are designated by superscripts.

Example:

$\underline{PGM}_1^1$, $\underline{PGM}_1^2$, $\underline{PGM}_1^3$, $\underline{PGM}_1^4$ etc., for alleles at the $\underline{PGM}_1$ locus.

The superscripts may be numerical or alphabetical. In rare cases, + and - signs which have been used extensively in the past may be retained.

Example:

$\underline{Gd}^B$, $\underline{Gd}^A$, $\underline{Gd}^{A-}$ for the three common alleles at the glucose-6-phosphate dehydrogenase locus in Black populations.

In other cases, place names are best used as the allele superscript to avoid confusion.

Example:

$\underline{Gd}^{Mediterranean}$, $\underline{Gd}^{Canton}$, $\underline{Gd}^{Athens}$, $\underline{Gd}^{Seattle}$

(Abbreviation of the place name may be more convenient.)

So-called "null" or "silent" alleles with little or no associated enzyme activity are best designated by the superscript 0 (i.e. zero), although the letter s may be retained because of common usage.

Examples:

$\underline{PGM}_1^0$, $\underline{E}_1^S$ ("silent" allele of the serum cholinesterase first locus)

When heterogeneity between "null" alleles can be demonstrated, the allele designation should be qualified, as by a place name.

Example:

$\underline{ADA}^0$ Calcutta

C. Examples of Genotypes

The following are some typical examples of genotypes written in accordance with the above recommendations and section D (below).

1. Heterozygote for the two common alleles at the ADA locus:

   $\underline{ADA^1ADA^2}$ (or $\underline{ADA^1}/\underline{ADA^2}$)

2. Heterozygotes for one or the other of these common ADA alleles and a "null" allele not separable from other "null" alleles at this locus:

   $\underline{ADA^1ADA^0}$ and $\underline{ADA^2ADA^0}$ (or $\underline{ADA^1}/\underline{ADA^0}$ and $\underline{ADA^2}/\underline{ADA^0}$) .

3. Genotype of an individual heterozygous for the two common alleles of $\underline{PGM_1}$, homozygous for the common allele of $\underline{PGM_2}$ and heterozygous for the two common alleles of $\underline{PGM_3}$ (3 unlinked loci):

   $\underline{PGM_1^1}/\underline{PGM_1^2}$, $\underline{PGM_2^1}/\underline{PGM_2^1}$, $\underline{PGM_3^1}/\underline{PGM_3^2}$

   or

   $$\frac{\underline{PGM_1^1}}{\underline{PGM_1^2}} \quad \frac{\underline{PGM_2^1}}{\underline{PGM_2^1}} \quad \frac{\underline{PGM_3^1}}{\underline{PGM_3^2}}$$

D. Linkage and Phase

A slash, either horizontal or semivertical (— or /) separating alleles, implies chromosomal location. The slash may be omitted in designating the genotype at a single locus. However, if two or more loci are involved, a horizontal line is recommended, particularly if the loci are syntenic.

1. Non-syntenic loci may be designated either by an interrupted horizontal line or by individual slashes and separation by commas.

Example:

$$\frac{\underline{ADA^1}}{\underline{ADA^2}} \quad \frac{\underline{PGM_1^1}}{\underline{PGM_2^2}} \quad \text{or} \quad \underline{ADA^1}/\underline{ADA^2}, \quad \underline{PGM_1^1}/\underline{PGM_2^2}$$

2. When the loci are in the same linkage group and the phase is known, the horizontal line is continuous.

Example:

$$\frac{\underline{AMY_1^A} \quad \underline{AMY_2^B}}{\underline{AMY_1^B} \quad \underline{AMY_2^A}}$$

(i.e. $\underline{AMY_1^A}$ and $\underline{AMY_2^B}$ are in cis position, as are their alleles)

3. When the loci are in the same linkage group but the phase is not known, a semicolon is used.

Example:

$$\frac{\underline{AMY_1^A}}{\underline{AMY_1^B}} ; \frac{\underline{AMY_2^A}}{\underline{AMY_2^B}}$$

4. To designate loci which are syntenic but not in the same linkage group, a colon is used.

Example:

$$\frac{\underline{AMY_2^A}}{\underline{AMY_2^B}} : \frac{\underline{PGM_1^1}}{\underline{PGM_1^2}}$$

II. Phenotypes

A. The phenotypic designation should have the same letters and subscripts as the locus (but not italicised or underlined), followed by the numerical, alphabetical or other symbol for the alleles, but not as superscripts. In the case of homozygotes for any allele or heterozygotes for a "null" allele, only one allele symbol is used.

Examples:

Genotype	Phenotype
$\underline{ADA^1}\underline{ADA^1}$	ADA 1
$\underline{ADA^1}\underline{ADA^2}$	ADA 2-1
$\underline{ADA^2}\underline{ADA^2}$	ADA 2
$\underline{ADA^1}\underline{ADA^0}$	ADA 1
$\underline{ADA^2}\underline{ADA^0}$	ADA 2
$\underline{PGM_1^1}/\underline{PGM_1^2}$, $\underline{PGM_2^1}/\underline{PGM_2^1}$, $\underline{PGM_3^1}/\underline{PGM_3^2}$	PGM_1 2-1, PGM_2 1, PGM_3 2-1

For hemizygotes, heterozygotes and homozygotes of the X-linked phosphoglycerate kinase alleles $\underline{PGK^1}$ and $\underline{PGK^2}$,

Genotype	Phenotype
$\underline{PGK^1}$	PGK 1
$\underline{PGK^2}$	PGK 2
$\underline{PGK^1}\underline{PGK^1}$	PGK 1
$\underline{PGK^1}\underline{PGK^2}$	PGK 2-1
$\underline{PGK^2}\underline{PGK^2}$	PGK 2

III. Isozyme Subunits

When two or more loci code for different polypeptide chains which occur together as subunits of single isozymes in a set of isozymes, it is useful to designate the subunit structure of the individual isozymes. Greek letters are convenient symbols for the polypeptide chains. A different letter can be used for the peptide product of each locus, by analogy with the α, β, γ and δ chains of hemoglobin. Whenever there are two or more alleles at a given locus coding for structurally different forms

of the same polypeptide, superscripts are incorporated which are the same as the superscripts used to designate the corresponding alleles.

Example:

The three loci of alcohol dehydrogenase, $\underline{ADH_1}$, $\underline{ADH_2}$ and $\underline{ADH_3}$ are thought to code for three different polypeptide chains: α, β and γ. There is evidence for two common alleles at the $\underline{ADH_2}$ locus: $\underline{ADH_2^1}$ and $\underline{ADH_2^2}$. These alleles code for polypeptides β^1 and β^2. There are also two common alleles at the $\underline{ADH_3}$ locus: $\underline{ADH_3^1}$ and $\underline{ADH_3^2}$, which code for polypeptides γ^1 and γ^2. All of the ADH isozymes are dimeric and the subunits interact with each other. In adult liver, all three loci are active. Thus, some of the isozymes are homodimers and some are heterodimers. The heteromeric isozymes contain polypeptides coded by alleles at either the same locus or at different loci. Thus, if an individual has the genotype

$$\underline{ADH_1^1ADH_1^1};\quad \underline{ADH_2^1ADH_2^1};\quad \underline{ADH_3^1ADH_3^2}$$

the phenotype is ADH_1 1, ADH_2 1, ADH_3 2-1

and in the electrophoretic pattern of a liver extract, there are ten isozymes with the following subunit structures:

$\alpha\alpha$	$\gamma^1\gamma^1$	$\alpha\gamma^1$	$\beta^1\gamma^1$
$\alpha\beta^1$	$\gamma^1\gamma^2$	$\alpha\gamma^2$	$\beta^1\gamma^2$
$\beta^1\beta^1$	$\gamma^2\gamma^2$		

In the following table, the enzyme name given is usually that recommended in 1972 by the Enzyme Commission.* When the E.C. name has not been used as the basis for the symbol, or if another name is much more familiar, the E.C. name is given first, and enclosed in brackets. (In a few instances the E.C. name is not given because it is so similar to the more familiar name.) The locus symbol given first is that recommended by this committee. Alternatives are also listed; these are based on systematic or obsolete names which can nearly always be found in the reference.* The computer symbols in the table are meant to be initial suggestions; they may require individual revision. The final column indicates that the given locus has been reported to be polymorphic in at least one large ethnic group.

*Enzyme Nomenclature: Recommendations (1972) of the International Union of Pure and Applied Chemistry and the International Union of Biochemistry. Published in 1973 by Elsevier (Amsterdam) and American Elsevier (New York).

Enzyme Name of phenotype	VAM number	E.C. No. (unless not applicable = NA)	Locus	Alternatives	chromosome assignment or linkage group	Computer Symbol	Polymorphic
Acid phosphatase-1		3.1.3.2	ACP_1		2 (P)	ACP-1	Yes
Acid phosphatase-2		3.1.3.2	ACP_2			ACP-2	
Acid phosphatase-3		3.1.3.2	ACP_3			ACP-3	
onitate hydratase]		4.2.1.3					
Aconitase (sol)		4.2.1.3	$ACON_S$	ACO_S, ACO_1		ACO-1	Yes
Aconitase (mito)		4.2.1.3	$ACON_M$	ACO_M, ACO_2		ACO-2	
Adenine phosphoribosyltransferase		2.4.2.7	APRT			APRT	
Adenosine deaminase		3.5.4.4	ADA			ADA	Yes
AMP deaminase		3.5.4.6	AMPDA			AMPDA	
Adenylate kinase-1		2.7.4.3	AK_1	AK-1		AK-1	Yes
Adenylate kinase-2		2.7.4.3	AK_2	AK-2		AK-2	
Alcohol dehydrogenase-1		1.1.1.1	ADH_1			ADH-1	
Alcohol dehydrogenase-2		1.1.1.1	ADH_2			ADH-2	Yes
Alcohol dehydrogenase-3		1.1.1.1	ADH_3			ADH-3	Yes
uctose-biphosphate aldolase]		4.1.2.13					
Aldolase-A		4.1.2.13	ALD_A	ALD_1		ALD-A	
Aldolase-B		4.1.2.13	ALD_B	ALD_2		ALD-B	
Aldolase-C		4.1.2.13	ALD_C	ALD_3		ALD-C	
Alkaline phosphatase (placental)		3.1.3.1	PL			ALPL	Yes
α Amylase (salivary)		3.2.1.1	AMY_1	AMY_S		AMY-1	Yes
α Amylase (pancreatic)		3.2.1.1	AMY_2	AMY_P		AMY-2	Yes
Aryl sulfatase		3.1.6.1	ARS			ARS	
rbonate dehydratase]		4.2.1.1					
Carbonic anhydrase-1		4.2.1.1	CA_1	CA_B, CA_I		CA-1	
Carbonic anhydrase-2		4.2.1.1	CA_2	CA_C, CA_{II}		CA-2	Yes
Catalase		1.11.1.6	CAT			CAT	
Cholinesterase (serum) -1		3.1.1.8	E_1			E-1	Yes

Enzyme Name	E.C. No.	Locus	Alternatives	Computer Symbol	Polymorphic
Cholinesterase (serum) -2	3.1.1.8	E_2		E-2	Yes
Citrate synthase	4.1.3.7	CS		CITSY	
Cytidine deaminase	3.5.4.5	CDA		CDA	Yes
Cytochrome b_5 reductase]	1.6.2.2				
Diaphorase (NADH)	1.6.2.2	DIA_1	DIA-A, Dia A	DIA-A	
Diaphorase (NADPH)	1.6.*.*	DIA_2	DIA-B, Dia B	DIA-B	Yes
2,3 Diphosphoglyceromutase	2.7.5.4	DPGM		DPGM	
Enolase-1	4.2.1.11	ENO_1	PPH_1	ENO-1	
Enolase-2	4.2.1.11	ENO_2	PPH_2	ENO-2	
Carboxylesterase]	3.1.1.1				
Esterase A_4	3.1.1.1	ESA_4	$Es\text{-}A_4$	ESA-4	
Esterase D	3.1.1.1	ESD		ESD	Yes
α-L-fucosidase	3.2.1.51	αFUC		A-FUC	Yes
Fumarate hydratase	4.2.1.2	FH		FUMH	
Galactokinase	2.7.1.6	GALK	GK, GAK	GALK	
Hexose-1-phosphate uridylyltransferase]	2.7.7.12				
Galactose-1-phosphate uridylyltransferase	2.7.7.12	GALT	Gt, Gal-1-PUT	GAPUT	Yes
α Galactosidase	3.2.1.22	αGAL		A-GAL	
Glucose-6-phosphate dehydrogenase	1.1.1.49	Gd	G6PD	G6PDH	Yes
Glucose phosphate isomerase	5.3.1.9	GPI	PHI	GPI	
α Glucosidase	3.2.1.20	αGLU		A-GLU	Yes
Aspartate aminotransferase]	2.6.1.1				
Glutamic-oxaloacetic transaminase (sol)	2.6.1.1	GOT_S	GOT-1, GOT_1	GOT-1	
Glutamic-oxaloacetic transaminase (mito)	2.6.1.1	GOT_M	GOT-2, GOT_2	GOT-2	Yes
Alanine aminotransferase]	2.6.1.2				
Glutamic-pyruvic transaminase	2.6.1.2	GPT		GPT	Yes
Glutathione peroxidase	1.11.1.9	GPX		GPX	

Enzyme Name	E.C. No.	Locus	Alternatives	Computer Symbol	Polymorphic
Glutathione reductase	1.6.4.2	GSR		GSR	Yes
Glyceraldehyde-3-phosphate dehydrogenase	1.2.1.12	GAPDH	GAPD	GAPDH	
Glycerol-3-phosphate dehydrogenase-1	1.1.1.8	GPD_1		GPD-1	
Glycerol-3-phosphate dehydrogenase-2	1.1.1.8	GPD_2		GPD-2	
Lactoyl-glutathione lyase]	4.4.1.5				
Glyoxylase I	4.4.1.5	GLO	GLY-1, Glx-1	GX-1	Yes
Hydroxyacylglutathione hydrolase]	3.1.2.6				
Glyoxylase II	3.1.2.6	HAGH	GLY-2, Glx-2	GX-2	
Guanylate kinase-1	2.7.4.8	GUK_1	GuK_1, $GMPK_1$	GMPK-1	
Guanylate kinase-2	2.7.4.8	GUK_2	GuK_2, $GMPK_2$	GMPK-2	
Guanylate kinase-3	2.7.4.8	GUK_3	GuK_3, $GMPK_3$	GMPK-3	
Hexokinase-1	2.7.1.1	HK_1	HK_I	HK-1	
Hexokinase-2	2.7.1.1	HK_2	HK_{II}	HK-2	
Hexokinase-3	2.7.1.1	HK_3	HK_{III}	HK-3	Yes
Hexokinase-4	2.7.1.1	HK_4	HK_{IV}	HK-4	
-N-acetylglucosaminidase]	3.2.1.30				
Hexosaminidase-A	3.2.1.30	HEX_A	$NAGA_A$	HEX-A	
Hexosaminidase-B	3.2.1.30	HEX_B	$NAGA_B$	HEX-B	
Hexosaminidase-C	3.2.1.30	HEX_C	$NAGA_C$	HEX-C	
Hypoxanthine phosphoribosyltransferase	2.4.2.8	HPRT	HGPRT	HGPRT	
ucleosidetriphosphate pyrophosphatase]	3.6.1.19				
Inosine triphosphatase	3.6.1.19	ITP		ITP	
Isocitrate dehydrogenase (sol)	1.1.1.42	ICD_S	IDH_S, IDH-1	ICDH-1	
Isocitrate dehydrogenase (mito)	1.1.1.42	ICD_M	IDH_M, IDH-2	ICDH-2	
Lactate dehydrogenase A	1.1.1.27	LDH_A	LDH-A	LDH-A	
Lactate dehydrogenase B	1.1.1.27	LDH_B	LDH-B	LDH-B	
Lactate dehydrogenase C	1.1.1.27	LDH_C	LDH-C	LDH-C	

Enzyme Name	E.C. No.	Locus	Alternatives	Computer Symbol	Polymorphic
Lecithin acyltransferase	2.3.1.43	LCAT		LCAT	
Malate dehydrogenase, NAD (sol)	1.1.1.37	MDH_S	MOR_S, MOR-1, MDH-1	MDH-1	
Malate dehydrogenase, NAD (mito)	1.1.1.37	MDH_M	MOR_M, MOR-2, MDH-2	MDH-2	
Malate dehydrogenase, decarb., $NADP^+$]	1.1.1.40				
Malic enzyme (sol)	1.1.1.40	ME_S	MOD_S, MOD-1, ME-1	ME-1	
Malic enzyme (mito)	1.1.1.40	ME_M	MOD_M, MOD-2, ME-2	ME-2	Yes
Mannose phosphate isomerase	5.3.1.8	MPI		MANPI	
Purine nucleoside phosphorylase]	2.4.2.1				
Nucleoside phosphorylase	2.4.2.1	NP		PURNP	
Pepsinogen (Pepsin)	3.4.23.*	Pg	Pg-5	PEPSG	Yes
Peptidase A	3.4.11.*	PEPA		PEP A	Yes
Peptidase B	3.4.11.*	PEPB		PEP B	
Peptidase C	3.4.11.*	PEPC		PEP C	Yes
Proline dipeptidase]	3.4.13.9				
Peptidase D	3.4.13.9	PEPD		PEP D	Yes
6-Phosphofructokinase A	2.7.1.11	PFK_A		PFK-A	
6-Phosphofructokinase B	2.7.1.11	PFK_B		PFK-B	
6-Phosphofructokinase C	2.7.1.11	PFK_C		PFK-C	
Phosphoglucomutase 1	2.7.5.1	PGM_1		PGM-1	Yes
Phosphoglucomutase 2	2.7.5.1	PGM_2		PGM-2	Yes
Phosphoglucomutase 3	2.7.5.1	PGM_3		PGM-3	Yes
Phosphogluconate dehydrogenase	1.1.1.44	PGD	6PGD	6PGD	Yes
Phosphoglycerate kinase	2.7.2.3	PGK		PGAK	
Phosphoglyceromutase (muscle)	2.7.5.3	$PGAM_M$		PGAM-M	
Phosphoglyceromutase (brain)	2.7.5.3	$PGAM_B$		PGAM-B	
Pyrophosphatase (inorganic)	3.6.1.1	PP		PP	
Pyruvate kinase (L)	2.7.1.40	PK_L	PK_I, PK-1	PK-L	

Enzyme Name	E.C. No.	Locus	Alternatives	Computer Symbol	Polymorphic
Pyruvate kinase (M1)	2.7.1.40	PK_{M1}		PK-M1	
Pyruvate kinase (M2)	2.7.1.40	PK_{M2}	PK_{III}, PK-3	PK-M2	
Ribosephosphate pyrophosphokinase	2.7.6.1	RPPK		RPPPK	
Iditol dehydrogenase]	1.1.1.14				
Sorbitol dehydrogenase	1.1.1.14	SORDH	SORD, SDH	SORDH	
Superoxide dismutase (sol)	1.15.1.1	SOD_S	IPO-A, SOD-A, SOD-1	SOD-1	
Superoxide dismutase (mito)	1.15.1.1	SOD_M	IPO-B, SOD-B, SOD-2	SOD-2	
Thymidine kinase	2.7.1.75	TK		TK	
Transaldolase	2.2.1.2	TRALD		TRALD	
Transketolase	2.2.1.1	TRKT		TRKT	
Triosephosphate isomerase	5.3.1.1	TPI		TPI	
ucose-1-phosphate uridylyltransferase]	2.7.7.9				
UDP Glucose pyrophosphorylase	2.7.7.9	UGPP		UGPP	
cleoside diphosphate kinase]	2.7.4.6				
Uridine diphosphate kinase	2.7.4.6	UDPK		UDPK	
Uridine monophosphate kinase	2.7.4.*	UMPK		UMPK	Yes
nucleotidase]	3.1.3.5				
Uridine monophosphatase	3.1.3.5	UMPH		UMPH	